<110> GAO, Guangxia
       LIU, Shufeng

<120> The Recombinant Murine Leukemia Virus Reverse Transcriptase,
       the Genes Encoding and the Method for Expressing it

<130> Q95957

<140> US 10/585,863
<141> 2006-07-11

<150> PCT/CN04/00039
<151> 2004-01-13

<160> 8

<210> 1
<211> 7488
<212> DNA
<213> artificial sequence

<220>
<223> plasmid DNA

<400> 1
ccgacaccat cgaatggtgc aaaacctttc gcggtatggc atgatagcgc ccggaagaga   60
gtcaattcag ggtggtgaat gtgaaaccag taacgttata cgatgtcgca gagtatgccg 120
gtgtctctta tcagaccgtt tcccgcgtgg tgaaccaggc cagccacgtt tctgcgaaaa 180
cgcgggaaaa agtggaagcg gcgatggcgg agctgaatta cattcccaac cgcgtggcac 240
aacaactggc gggcaaacag tcgttgctga ttggcgttgc cacctccagt ctggccctgc 300
acgcgccgtc gcaaattgtc gcggcgatta aatctcgcgc cgatcaactg ggtgccagcg 360
tggtggtgtc gatggtagaa cgaagcggcg tcgaagcctg taaagcggcg gtgcacaatc 420
ttctcgcgca acgcgtcagt gggctgatca ttaactatcc gctggatgac caggatgcca 480
ttgctgtgga agctgcctgc actaatgttc cggcgttatt tcttgatgtc tctgaccaga 540
cacccatcaa cagtattatt ttctcccatg aagacggtac gcgactgggc gtggagcatc 600
tggtcgcatt gggtcaccag caaatcgcgc tgttagcggg cccattaagt tctgtctcgg 660
cgcgtctgcg tctggctggc tggcataaat atctcactcg caatcaaatt cagccgatag 720
cggaacggga aggcgactgg agtgccatgt ccggttttca acaaaccatg caaatgctga 780
atgagggcat cgttcccact gcgatgctgg ttgccaacga tcagatggcg ctgggcgcaa 840
tgcgcgccat taccgagtcc gggctgcgcg ttggtgcgga tatctcggta gtgggatacg 900
acgataccga agacagctca tgttatatcc cgccgttaac caccatcaaa caggattttc 960
gcctgctggg gcaaaccagc gtggaccgct tgctgcaact ctctcagggc caggcggtga 1020
agggcaatca gctgttgccc gtctcactgg tgaaaagaaa aaccaccctg gcgcccaata 1080
cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca cgacaggttt 1140
cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag 1200
gcacaattct catgtttgac agcttatcat cgactgcacg gtgcaccaat gcttctggcg 1260
tcaggcagcc atcggaagct gtggtatggc tgtgcaggtc gtaaatcact gcataattcg 1320
tgtcgctcaa ggcgcactcc cgttctggat aatgtttttt gcgccgacat cataacggtt 1380
ctggcaaata tctgcgaaatg agctgttgac aattaatcat cggctcgtat aatgtgtgga 1440
attgtgagcg gataacaatt tgaattctta agatttgtga ggggataaca atttcacaca 1500
ggaaacagaa tatgacccta aatatagagg atgcatcatg gctacaaaag acctcaaaag 1560
agccagatgt ttctctaggg tccacatggc tgtctgattt tcctcaggcc tgggcggaaa 1620
ccggggggcat gggactggca gttcgccaag ctcctctgat catacctctg aaagcaacct 1680
ctaccccccgt gtccataaaa caatacccca ttcacaaga agcagactg gggatcaagc 1740
cccacataca gagactgttg gacgcgggaa ttctggtacc ctgccagtcc ccctggaaca 1800
cgcccctgct acccgttaag aaaccaggga ctaatgatta taggcctgtc caggatctga 1860
gagaagtcaa caagcgggtg gaagacatcc accccgcct gcccaaccct tacaacctct 1920
tgagcgggct cccaccgtcc caccagtggt acactgtgct tgatttaaag gatgcctttt 1980
tctgcctgag actccacccc accagtcctc gtcttcgc ctttgagtgg agagatccag 2040
agatgggaat ctcaggacaa ttgacctgga ccagactccc acaggggttc aaaaacagtc 2100
ccaccctgtt tgatgaggca ctgcacagag acctagcaga cttccggga tc cagcacccag 2160

```
acttgatcct gctacagtac gtggatgact tactgctggc cgccacttct gagctagact 2220
gccaacaagg tactcgggcc ctgttacaaa ccctagggaa cctcgggtat cgggcctcgg 2280
ccaagaaagc ccaaatttgc cagaaacagg tcaagtatct ggggtatctt ctaaaagagg 2340
gtcagagatg gctgactgag gccagaaaag agactgtgat ggggcagcct actccgaaga 2400
ccctcgaca actaagggag ttcctaggga cgcagagact ctgtcgcctc tggatccctg 2460
ggtttgcaga aatggcagcc ccctgtacc ctctcaccaa aacggggact ctgtttaatt 2520
ggggcccaga ccaacaaaag gcctatcaag aaatcaagca agtcttctca actgcccccag 2580
ccctgggtt gccagatttg actaagccct tgtaactctt tgtcgacgag aagcagggct 2640
acgccaaagg tgtcctaacg caaaaactgg gaccttggcg tcggccggtg gcctacctgt 2700
ccaaaaagct agacccagta gcagctgggt ggccccccttg cctacggatg gtagcagcca 2760
ttgccgtact gacaaaggat gcaggcaagc taaccatggg acagccacta gtcattctgg 2820
cccccccatgc agtagaggca ctagtcaaac aaccccccga ccgctggctt tccaacgccc 2880
ggatgactca ctatcaggcc ttgcttttgg acacggaccg ggtccagttc ggaccggtgg 2940
tagccctgaa cccggctacg ctgctcccac tgcctgagga agggctgcaa cacaactgcc 3000
ttgatatcct ggccgaagcc cacggaaccc gacccgacct aacggaccag ccgctcccag 3060
acgccgacca cacctggtac acgaatggaa gcagtctctt acaagaggga cagcgtaagg 3120
cgggagctgc ggtgaccacc gagaccgagg taatctgggc taaagccctg ccagccggga 3180
catccgctca gcgggctgaa ctgatagcac tcacccaggc cctaaaagtg gcagaaggta 3240
agaagctaaa tgtttatact gatagccgtt atgctttgc tactgcccat atccatggag 3300
aaatatacag aaggcgtggg ttgctcacat cagaagacaa agagatcaaa aataagacg 3360
agatcttggc cctactaaaa gcctctttctc tgcccaaaag acttagcata atccattgtc 3420
caggacata aaaggacac agcgccgagg ctagaggcaa ccggatggct gaccaagcgg 3480
cccgaaaggc agccatcaca gagactccag acacctctac cctcctccat caccatcacc 3540
atcactagtc tagagtcgac ctgcaggcaa gcttggcacc ggccgtcgtt ttacaacgtc 3600
gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat cccccttttcg 3660
ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc 3720
tgaatggcga atggcgcctt ggctgtttg gcggatgaga taagatttc agcctgatca 3780
agattaaatc agaacgcaga agcggtctga taaaacagaa tttgcctggc ggcagtagcg 3840
cggtggtccc acctgacccc atgccgaact cagaagtgaa acgccgtagc gccgatggta 3900
gtgtgggggtc tccccatgcg agagtaggga actgccaggc atcaaataaa acgaaaggct 3960
cagtcgaaag actgggcctt tcgttttatc tgttgtttgt cggtgaacgc tctcctgagt 4020
aggacaaatc cgccgggagc ggatttgaac gttgcgaagc aacggcccgg agggtggcgg 4080
gcaggacgcc cgccataaac tgccaggcat caaattaagc agaaggccat cctgacggat 4140
ggcctttttg cgtttctaca aactctttt gtttatttt ctaaatacat tcaaatatgt 4200
atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta 4260
tgagtattca acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg 4320
tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac 4380
gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg 4440
aagaacgttc tccaatgatg agcacttta aagttctgct atgtggcgcg gtattatccc 4500
gtgttgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg 4560
ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaaattat 4620
gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg 4680
gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg 4740
atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc 4800
ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt 4860
cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct 4920
cggcccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc 4980
gcggtatcat tgcagcactg gggccagatg gtaagccctc cgtatctgac gggcagcaat 5040
cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct 5100
cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt 5220
taccccggtt gataatcaga aaagccccaa aaacaggaag attgtataag caaatattta 5220
aattgtaaac gttaaatatt tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt 5280
ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agcccgagat 5340
aggggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa 5400
cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac catcacccaa 5460
atcaagtttt ttggggtcga ggtgccgtaa agcactaaat cggaaccta aagggagccc 5520
ccgattaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc 5580
gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc acgctgcgcg taaccaccac 5640
acccgccgcg cttaatgcgc cgctacaggg cgcgtaaaag gctaggtg aagatccttt 5700
ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc 5760
ccgtagaaaa gatcaaagga tcttcttgag atcttttttt tctgcgcgta atctgctgct 5820
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa 5880
ctcttttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttcctag 5940
```

```
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc 6000
tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg 6060
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 6120
cacagcccag cttggagcga acgacctaca ccgaactgag ataccacaca cgtgagctat 6180
gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtattcggta acgccgcaggg 6240
tcggaacagg agacgcgacg agggagctc cagggggaa cgcctggtat cttatagtc 6300
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 6360
ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc 6420
cttttgctca catgttctt cctgcgttat cccctgattc tgtggataac cgtattaccg 6480
cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga 6540
gcgaggaagc ggaagagcgc ctgatgcggt attttctcct tacgcatctg tgcggtattt 6600
cacaccgcat atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt 6660
atacactccg ctatcgctac gtgactgggt catggctgcg ccccgacacc cgccaacacc 6720
cgctgacgcg ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac 6780
cgtctccggg agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca 6840
gctgcggtaa agctcatcag cgtggtcgtg cagcgattca cagatgtctg cctgttcatc 6900
cgcgtccagc tcgttgagtt tctccagaag cgttaatgtc tggcttctga taaagcgggc 6960
catgttaagg gcggtttttt cctgtttggt cacttgatgc ctccgtgtaa gggggaattt 7020
ctgttcatgg gggtaatgat accgatgaaa cgagagagga tgctcacgat acgggttact 7080
gatgatgaac atgcccggtt actggaacgt tgtgagggta aacaactggc ggtatggatg 7140
cggcgggacc agagaaaaat cactcagggt caccggctggc gcttcgttaa tacagatgta 7200
ggtgttccax agggtagcca gcagcatcct gcgatgcaga tgcgggaacat aatggtgcag 7260
ggcgctgact tccgcgtttc cagactttac gaaacacgga aaccgaagac cattcatgtt 7320
gttgctcagg tcgcagacgt tttgcagcag cagtcgcttc acgttcgctc gcgtatcgat 7380
gattcattct gctaaccagt aaggcaaccc cgccagccta gccgggtcct caacgacagg 7440
agcacgatca tgcgcacccg tggccaggac ccaacgctgc ccgaaatt              7488
```

```
<210>  2
<211>  671
<212>  PRT
<213>  artificial sequence

<220>
<223>  reverse transcriptase

<400>  2
Thr Leu Asn Ile Glu Asp Glu His Arg Leu His Glu Thr Ser Lys Glu
1               5                   10                  15
Pro Asp Val Ser Leu Gly Ser Thr Trp Leu Ser Asp Phe Pro Gln Ala
                20                  25                  30
Trp Ala Glu Thr Gly Gly Met Gly Leu Ala Val Arg Gln Ala Pro Leu
            35                  40                  45
Ile Ile Pro Leu Lys Ala Thr Ser Thr Pro Val Ser Ile Lys Gln Tyr
        50                  55                  60
Pro Met Ser Gln Glu Ala Arg Leu Gly Ile Lys Pro His Ile Gln Arg
65                  70                  75                  80
Leu Leu Asp Ala Gly Ile Leu Val Pro Cys Gln Ser Pro Trp Asn Thr
                85                  90                  95
Pro Leu Leu Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr Arg Pro Val
                100                 105                 110
Gln Asp Leu Arg Glu Val Asn Lys Arg Val Glu Asp Ile His Pro Thr
            115                 120                 125
Val Pro Asn Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro Ser His Gln
        130                 135                 140
Trp Tyr Thr Val Leu Asp Leu Lys Asp Ala Phe Phe Cys Leu Arg Leu
145                 150                 155                 160
His Pro Thr Ser Gln Pro Leu Phe Ala Phe Glu Trp Arg Asp Pro Glu
                165                 170                 175
Met Gly Ile Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro Gln Gly Phe
                180                 185                 190
Lys Asn Ser Pro Thr Leu Phe Asp Glu Ala Leu His Arg Asp Leu Ala
            195                 200                 205
Asp Phe Arg Ile Gln His Pro Asp Leu Ile Leu Leu Gln Tyr Val Asp
```

```
    210                         215                         220
Asp Leu Leu Leu Ala Ala Thr Ser Glu Leu Asp Cys Gln Gln Gly Thr
225                         230                         235                         240
Arg Ala Leu Leu Gln Thr Leu Gly Asn Leu Gly Tyr Arg Ala Ser Ala
                    245                         250                         255
Lys Lys Ala Gln Ile Cys Gln Leu Val Lys Tyr Leu Ser Gly Tyr Leu
                260                         265                         270
Leu Lys Gly Gly Gln Arg Trp Leu Thr Glu Ala Arg Lys Glu Thr Val
            275                         280                         285
Met Gly Gln Pro Thr Pro Lys Thr Pro Arg Gln Leu Arg Glu Phe Leu
        290                         295                         300
Gly Thr Ala Gly Phe Cys Arg Leu Trp Ile Pro Gly Phe Ala Glu Met
305                         310                         315                         320
Ala Ala Pro Leu Tyr Pro Leu Thr Lys Thr Gly Thr Leu Phe Asn Trp
                    325                         330                         335
Gly Pro Asp Gln Gln Lys Ala Tyr Gln Gln Ile Lys Gln Ala Leu Leu
                340                         345                         350
Thr Ala Pro Ala Leu Gly Leu Pro Asp Leu Thr Lys Pro Phe Glu Leu
            355                         360                         365
Phe Val Asp Glu Lys Gln Gly Tyr Ala Lys Gly Val Leu Thr Gln Lys
        370                         375                         380
Leu Gly Pro Trp Arg Arg Pro Val Ala Tyr Leu Ser Lys Lys Leu Asp
385                         390                         395                         400
Pro Val Ala Ala Gly Trp Pro Pro Cys Leu Arg Met Val Ala Ala Ile
                    405                         410                         415
Ala Val Leu Thr Lys Asp Ala Gly Lys Leu Thr Met Gly Gln Pro Leu
                420                         425                         430
Val Ile Leu Ala Pro His Ala Val Glu Ala Leu Val Lys Gln Pro Pro
            435                         440                         445
Asp Arg Trp Leu Ser Asn Ala Arg Met Thr His Tyr Gln Ala Leu Leu
        450                         455                         460
Leu Asp Thr Asp Arg Val Gln Phe Gly Pro Val Val Ala Leu Asn Pro
465                         470                         475                         480
Ala Thr Leu Leu Pro Leu Pro Glu Glu Gly Leu Gln His Asn Cys Leu
                    485                         490                         495
Asp Ile Leu Ala Glu Ala His Gly Thr Arg Pro Asp Leu Thr Asp Gln
                500                         505                         510
Pro Leu Pro Asp Ala Asp His Thr Trp Tyr Thr Asn Gly Ser Ser Leu
            515                         520                         525
Leu Gln Glu Gly Gln Arg Lys Ala Gly Ala Ala Val Thr Thr Glu Thr
        530                         535                         540
Glu Val Ile Trp Ala Lys Ala Leu Pro Ala Gly Thr Ser Ala Gln Arg
545                         550                         555                         560
Ala Glu Leu Ile Ala Leu Thr Gln Ala Leu Lys Met Ala Glu Gly Lys
                    565                         570                         575
Lys Leu Asn Val Tyr Thr Asp Ser Arg Tyr Ala Phe Ala Thr Ala His
                580                         585                         590
Ile His Gly Glu Ile Tyr Arg Arg Arg Gly Leu Leu Thr Ser Glu Gly
            595                         600                         605
Lys Glu Ile Lys Asn Lys Asp Glu Ile Leu Ala Leu Leu Lys Ala Leu
        610                         615                         620
Phe Leu Pro Lys Arg Leu Ser Ile Ile His Cys Pro Gly His Gln Lys
625                         630                         635                         640
Gly His Ser Ala Glu Ala Arg Gly Asn Arg Met Ala Asp Gln Ala Ala
                    645                         650                         655
Arg Lys Ala Ala Ile Thr Glu Thr Pro Asp Thr Ser Thr Leu Leu
                660                         665                         670
```

<210> 3
<211> 7488
<212> DNA
<213> artificial sequence

```
<220>
<223>  plasmid DNA

<400> 3
ccgacaccat cgaatggtgc aaaaccttc gcggtatggc atgatagcgc ccggaagaga      60
gtcaattcag ggtggtgaat gtgaaaccag taacgttata cgatgtcgca gagtatgccg     120
gtgtctctta tcagacgcgt tcccgcgtgg tgaaccaggc cagccacgtt tctgcgaaaa     180
cgcgggaaaa agtggaagcg gcgatggcgg agctgaatta cattcccaac cgcgtggccac    240
aacaactggc gggcaaacag tcgttgctga ttggcgttgc cacctccagt ctggccctgc     300
acgcgccgtc gcaaattgtc gcggcgatta aatctcgcgc cgatcaactg ggtgccagcg     360
tggtggtgtc gatggtagaa cgaagcggcg tcgaagcctg taaagcggcg gtgcacaatc     420
ttctcgcgca acgcgtcagt gggctgatca ttaactatcc gctggatgac caggatgcca     480
ttgctgtgga agctgcctgc actaatgttc cggcgttatt tcttgatgtc tctgaccagca    540
cacccatcaa cagtattatt ttctcccatg aaagcgtgg gcgcatggggc gtggagcatc     600
tggtgcgatt gggtcaccag caaatcgcgc tgttagcggg cccattaagt tctgtctcgg     660
cgcgtctgcg tctggctggc tggcataaat atctcactcg caatcaaatt cagccgatag     720
cggaacggga aggcgactgg agtgccatgt ccggttttca acaaaccatg caaatgctga     780
atggacgcat cgttccccact gcgatggcgg agctgaatta cattcccaac cgcgtggccac   840
tgcgcgccat taccgagtcc gggctgcgcg ttggtgcgga tatctcggta gtgggatacg     900
acgataccga agacgatgct gttgtatatc caccatcaaa caggattttc                960
gcctgctggg gcaaaccagc gtggaccgct tgctgcaact ctctcagggc caggcggtga    1020
agggcaatca gctgttgccc gtctcactgg tgaaaagaaa aaccaccctg gcgcccaata    1080
cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca cgacaggtttt  1140
cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag    1200
gcacaatctt catgtttgac agcttatcat cgactgcacg gtgcaccaat gcttctggcg    1260
tcaggcagcc atcggaagct gtggtatggc tgtgcaggtc gtaaatcact gcataattcg    1320
tgtcgctcaa ggcgcactcc cgttctggat aatgttttt gcgccgacat cataacggtt     1380
ctggcaaata ttctgaaatg agctgttgac aattaatcat cggctcgtat aatgtgtgga    1440
attgtgagcg gataacaatt tgaattctta agatttgtga ggggataaca atttcacaca    1500
ggaaacagaa tatgaccta aatatagaag atgagcatcg gctacatgag acctcaaaag     1560
agccagatgt ttctctaggg tccacatggc tgtctgattt tcctcaggcc tgggcggaaa    1620
ccgggggcat gggactggca gttcgccaag ctcctctgat cataccctg aaagcaacct     1680
ctacccccgt gtccataaaa caatacccca tgtcacaaga agccagactg gggatcaagc    1740
cccacataca gagactgttg gacaacggga tcctggtacc ctgccagtcc ccctggaaca    1800
cgccccctgct acccgttaag aaaccagggga ctaatgatta taggcctgtc caggatctga   1860
gagaagtcaa caagcgggtg gaagacatcc atcccaccgt gcccaaccct tacaaccttt    1920
tgagcgggct cccaccgtcc caccagtgat acactgtgct tgatttaaag gatgcctttt    1980
tctgcctgag actccaccc accagtcagc ctctcttcgc ctttgagtgg agagatccag    2040
agatgggaat ctcaggacaa ttgacctgga ccagctccc acagggtttc aaaaacagtc    2100
ccaccctgtt tgatgaggca ctgcacagag acctagcaga cttccggatc cagcacccag   2160
acttgatcct gctacagtac gtggatgact tactgctggc cgccacttct gagctagact    2220
gccaacaagg tactcgggcc ctgttacaaa ccctaagggaa cctcgggtat cgggcctcgg    2280
ccaagaaagc ccaaatttgc cagaaacagg tcaagtatct ggggtatctt ctaaaagagg    2340
gtcagagatg gctgactgag gccagaaaag agactgtgat gggcaacct actccgaaaa     2400
cccctcgaca actaagggag ttcctaggga cggcaggctt ctgtcgcctc tggatccctg    2460
ggtttgcaa aatggcgacc ccccttgtacc ctctcaccaa aacggggact ctgtttaatt    2520
ggggcccaga ccaacaaaag gcctatcaag aaatcaagca agtcttcta actgccccag     2580
ccctggggtt gccagatttg actaagccct ttgaactctt tgtcgacgag aagcagggct    2640
acgccaaagg tgtcctaacg caaaaactgg gaccttggcg tcggccggtg gcctacctgt    2700
ccaaaaagct agacccagca gcagctgggt ggccccctttg cctacggacg gtagcagcca    2760
ttgccgtact gacaaaggat gcaggcaagc taaccatggg acagccacta gtcattctgg    2820
ccccccatgc agtgaggca ctagtcaaac ccctcctgac tgtcgccaag gacctcaccg   2880
ggatgactca ctatcaggcc ttgtctttgg acacgagccg ggtctcaatc gggacctgtc   2940
tagcctgaa cccggctacg ctgtccccac tgctcctgag agggctgca cacaactgcc     3000
ttgatatcct ggccgaagcc caccgagggg acccctggaac aaccgggatc gccgccctt     3060
acgccgacca cacctggtac acgaatggaa gcagtctctt acaagagga cagcgtaagg    3120
cgggagctgc ggtgaccacc agaccgagag taatctgggc taaagccct gc cagccggga   3180
catccgctca gcgggctgaa ctgatacgac tcacccaggc cctaaagatg gcagaaggta    3240
agaagtaaa tgtttatact gatagccgtt atgctttttg tactgcccat atccatggaa    3300
aaatatacag aaggcgtggg ttgctcacat cagaaggcaa agagat caaa aataaaagcg   3360
agatcttgg cctactaaaa gccctcttc tccacaaag acttagcata atccattgtc      3420
caggacatca aaagggacac agcgccgagg ctagaggcaa ccggatggct gaccaagcgg    3480
cccgaaaggc agccatcaca gagatccag acacctctac cctcctccat caccatcacc    3540
```

Page 5

```
atcactagtc tagagtcgac ctgcaggcaa gcttggcact ggccgtcgtt ttacaacgtc    3600
gtgactgggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat cccccttttcg   3660
ccagctggcg taataggcga gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc    3720
tgaatggcga atggcgacgt ggctgtttg gcggatgaga taagattctc agcctgatac    3780
agattaaatc agaacgcaga agcggtctga taaaacagaa tttgcctggc ggcagtagcg    3840
cggtggtccc acctgaccc atgccgaact cagaagtgaa acgccgtagc gccgatggta    3900
gtgtggggtc tccccatgcg agagtaggga actgccaggc atcaaatraa acgaaaggct    3960
cagtcgaaag actgggcctt tcgttttatc tgttgtttgt cggtgaacgc tctcctgagt    4020
aggacaaatc cgccgggagc ggatttgaac gttgcgaagc aacggcccgg agggtggcgg    4080
gcaggacgcc cgccataaac tgccaggcat caaataagc agaaggccat cctgacggat    4140
ggccttttttg cgtttctaca aactctttt gtttatttt ctaaatacat tcaaatatgt     4200
atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta     4260
tgagtattca acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg     4320
tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac     4380
gagtgggtta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg     4440
aagaacgttc tccaatgatg agcactttaa aagttctgct atgtggcgcg gtattatccc     4500
gtgttgacgc cgggcaagaa caactcggtc gccgcataca ctattctcag aatgacttgg     4560
ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat     4620
gcagtgctgc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg     4680
gaggaccgga aggcgtaacc gcttttttgc acaacatggg ggatcatgta actcgccttg     4740
atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc     4800
ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt     4860
cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct     4920
cggccccttcc ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc     4980
gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca     5040
cgacgggag tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct     5100
cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt     5160
taccccggtt gataatcaga aaagccccaa aaacaggaag attgtataag caaatattta     5220
aattgtaaac gttaatattt tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt     5280
ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agcccgagat     5340
agggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa     5400
cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac catcacccaa     5460
atcaagtttt ttggggtcga ggtgccgtaa agcactaaat cggaacccta aagggagccc     5520
ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc     5580
gaaaggagcg ggcgctaagg cgctggcaag tgtagcggtc acgctgcgcg taaccaccac     5640
acccgccgcg cttaatgcgc cgctacaggg cgcgtaaaaa gatctaggtg aagatccttt     5700
ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc     5760
ccgtagaaaa gatcaaagga tcttcttgag atccttttttt tctgcgcgta atctgctgct     5820
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa     5880
ctcttttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag     5940
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc     6000
tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg     6060
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca     6120
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat     6180
gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg     6240
tcggaacagg agagcgcacg aggggagctTC cagggggaaa cgcctggtat ctttatagtc     6300
ctgtcgggtt tcgccacctc tgacttgagc gtcgatttt tgtgatgctcg tcaggggggc     6360
ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc     6420
cttttgctca catgttcttt cctgcgttat ccctgattc tgtggataac gtattaccg     6480
cctttgactg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga     6540
gcgaggaagc ggaagagcgc ctgatgcggt attttctcct tacgcatctg tgcggtattt     6600
cacaccgcat atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt     6660
atacactccg ctatcgctac gtgactggt catggtcgcg ccccgcaacc cgccacacc     6720
cgctgacgcg ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac     6780
cgtctccggg agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca     6840
gctgcggta agctcatcag cgtggtcgtg cagcgattca cagatctgtc cctgttcatc     6900
cgcgtccagc tcgttgagtt tctccagaag cgttaatgtc tggcttctga taaagcgggc     6960
catgttaagg gcggtttttt cctgtttggt cacttgatgc ctccgtgtaa ggggggaattt     7020
ctgttcatgg gggtaatgat accgatgaaa cgagagagga tgctcacgat acgggttact     7080
gatgatgaac atgcccggtt actgaacgt tgtgagggta aacaactggc ggtatggatg     7140
cggcgggacc agagaaaaat cactcaggt caatgccagc gcttcgttaa tacagatgta     7200
ggtgttccac aggtagcca gcagcatcct gcgatgcaga tccggaacat aatggtgcag     7260
ggcgctgact tccgcgtttc cagactttac gaaacacgga aaccgaagac cattcatgtt     7320
```

```
gttgctcagg tcgcagacgt tttgcagcag cagtcgcttc acgttcgctc gcgtatcggt    7380
gattcattct gctaaccagt aaggcaaccc cgccagccta gccgggtcct caacgacagg    7440
agcacgatca tgcgcacccg tggccaggac ccaacgctgc ccgaaatt                 7488
```

<210> 4
<211> 671
<212> PRT
<213> artificial sequence

<220>
<223> reverse transcriptase

<400> 4

```
Thr Leu Asn Ile Glu Asp Glu His Arg Leu His Glu Thr Ser Lys Glu
1               5                   10                  15
Pro Asp Val Ser Leu Gly Ser Thr Trp Leu Ser Asp Phe Pro Gln Ala
            20                  25                  30
Trp Ala Glu Thr Gly Gly Met Gly Leu Ala Val Arg Gln Ala Pro Leu
        35                  40                  45
Ile Ile Pro Leu Lys Ala Thr Ser Thr Pro Val Ser Ile Lys Gln Tyr
    50                  55                  60
Pro Met Ser Gln Glu Ala Arg Leu Gly Ile Lys Pro His Ile Gln Arg
65                  70                  75                  80
Leu Leu Asp Asn Gly Ile Leu Val Pro Cys Gln Ser Pro Trp Asn Thr
                85                  90                  95
Pro Leu Leu Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr Arg Pro Val
            100                 105                 110
Gln Asp Leu Arg Glu Val Asn Lys Arg Val Glu Asp Ile His Pro Thr
        115                 120                 125
Val Pro Asn Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro Ser His Gln
    130                 135                 140
Trp Tyr Thr Val Leu Asp Leu Lys Asp Ala Phe Phe Cys Leu Arg Leu
145                 150                 155                 160
His Pro Thr Ser Gln Pro Leu Phe Ala Phe Glu Trp Arg Asp Pro Glu
                165                 170                 175
Met Gly Ile Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro Gln Gly Phe
            180                 185                 190
Lys Asn Ser Pro Thr Leu Phe Asp Glu Ala Leu His Arg Asp Leu Ala
        195                 200                 205
Asp Phe Arg Ile Gln His Pro Asp Leu Ile Leu Leu Gln Tyr Val Asp
    210                 215                 220
Asp Leu Leu Leu Ala Ala Thr Ser Glu Leu Asp Cys Gln Gln Gly Thr
225                 230                 235                 240
Arg Ala Leu Leu Gln Thr Leu Gly Asn Leu Gly Tyr Arg Ala Ser Ala
                245                 250                 255
Lys Lys Ala Gln Ile Cys Gln Lys Gln Val Lys Tyr Leu Gly Tyr Leu
            260                 265                 270
Leu Lys Glu Gly Gln Arg Trp Leu Thr Glu Ala Arg Lys Glu Thr Val
        275                 280                 285
Met Gly Gln Pro Thr Pro Lys Thr Pro Arg Gln Leu Arg Glu Phe Leu
    290                 295                 300
Gly Thr Ala Gly Phe Cys Arg Leu Trp Ile Pro Gly Phe Ala Glu Met
305                 310                 315                 320
Ala Ala Pro Leu Tyr Pro Leu Thr Lys Thr Gly Thr Leu Phe Asn Trp
                325                 330                 335
Gly Pro Asp Gln Gln Lys Ala Tyr Gln Glu Ile Lys Gln Ala Leu Leu
            340                 345                 350
Thr Ala Pro Ala Leu Gly Leu Pro Asp Leu Thr Lys Pro Phe Glu Leu
        355                 360                 365
Phe Val Asp Glu Lys Gln Gly Tyr Ala Lys Gly Val Leu Thr Gln Lys
    370                 375                 380
Leu Gly Pro Trp Arg Arg Pro Val Ala Tyr Leu Ser Lys Lys Leu Asp
```

Page 7

```
385                     390                     395                     400
Pro Val Ala Ala Gly Trp Pro Pro Cys Leu Arg Met Val Ala Ala Ile
                    405                     410                     415
Ala Val Leu Thr Lys Asp Ala Gly Lys Leu Thr Met Gly Val Gln Pro Leu
            420                     425                     430
Val Ile Leu Ala Pro His Ala Val Glu Ala Leu Val Lys Gln Pro Pro
        435                     440                     445
Asp Arg Trp Leu Ser Asn Ala Arg Met Thr His Tyr Gln Ala Leu Leu
    450                     455                     460
Leu Asp Thr Asp Arg Gln Phe Gly Pro Val Val Ala Leu Asn Pro
465                     470                     475                     480
Ala Thr Leu Leu Pro Leu Pro Glu Gly Gly Leu Gln His Asn Cys Leu
                    485                     490                     495
Asp Ile Leu Ala Glu Ala His Gly Thr Arg Pro Asp Leu Thr Asp Gln
            500                     505                     510
Pro Leu Pro Asp Ala Asp His Thr Trp Tyr Thr Asn Gly Ser Ser Leu
        515                     520                     525
Leu Gln Glu Gly Gln Arg Lys Ala Gly Ala Ala Val Thr Thr Glu Thr
    530                     535                     540
Glu Val Ile Trp Ala Lys Ala Leu Pro Ala Gly Thr Ser Ala Gln Arg
545                     550                     555                     560
Ala Glu Leu Ile Ala Ala Thr Gln Ala Leu Lys Met Ala Glu Gly Lys
                    565                     570                     575
Lys Leu Asn Val Tyr Thr Asp Ser Arg Tyr Ala Phe Ala Thr Ala His
            580                     585                     590
Ile His Gly Glu Ile Tyr Arg Arg Arg Gly Leu Leu Thr Ser Glu Gly
        595                     600                     605
Lys Glu Ile Lys Asn Lys Asp Glu Ile Leu Ala Leu Leu Lys Ala Leu
    610                     615                     620
Phe Leu Pro Lys Arg Leu Ser Ile Ile His Cys Pro Gly His Gln Lys
625                     630                     635                     640
Gly His Ser Ala Glu Ala Arg Gly Asn Arg Met Ala Asp Gln Ala Ala
                    645                     650                     655
Arg Lys Ala Ala Ile Thr Glu Thr Pro Asp Thr Ser Thr Leu Leu
    660                     665                     670
```

```
<210>   5
<211>   17
<212>   DNA
<213>   artificial sequence

<220>
<223>   primer

<400>   5
gtggaattgt gagccga                                                      17

<210>   6
<211>   29
<212>   DNA
<213>   artificial sequence

<220>
<223>   primer

<400>   6
cggaattccc gcgtccaaca gtctctgta                                         29

<210>   7
<211>   17
<212>   DNA
<213>   artificial sequence
```

```
<220>
<223>  primer

<400>  7
tgggagtctg gtccagg                                              17

<210>  8
<211>  24
<212>  DNA
<213>  artificial sequence

<220>
<223>  primer

<400>  8
cggaattctg gtaccctgcc agtc                                     24
```